Control over the emergence of self-replicators in dynamic molecular networks

Leonetti, Giulia

Chapter 2

Solvent Composition Dictates Emergence in Dynamic Molecular Networks Containing Competing Replicators

This chapter has been published:

Leonetti, G.; Otto, S. *J. Am .Chem. Soc.* **137**, 2067-2072 (2015).

2.1. Introduction

The interplay between biological species and their environment has a central role in Darwinian evolution, since changes in external conditions force living species to evolve, increasing their fitness and population accordingly. Natural selection favors the offspring of a given species that is best adapted to the local environments. From an origin-of-life perspective, in order to understand how life could originate from inanimate matter,[1] it would be of great interest to investigate whether similar mechanisms occur at the molecular level, with fully synthetic systems.

The role of the environment on proteins was investigated with computer simulations by Earl and Deem in 2004.[2] The scientists reasoned that evolution of biological diversity should have been facilitated by the development of mechanisms that enabled the evolution of the species. Performing calculations on protein evolution, they found that the system selects those proteins having higher rates of evolvability instead of higher fitness if subjected to harsh environmental changes. Therefore, the selection of proteins for a specific environment is not optimal, since great fitness values for one environment would lead to a smaller evolvability when a change of conditions is faced by the system. Evolvability, therefore, is heavily influenced by "underlying rates of genetic change", that is the probability of errors in templating mechanisms (polymerases, recombination, transposition and so on). A more general computer simulation on the effects of the environment on evolutionary processes was reported by Alon and coworkers in 2007.[3] They defined evolution as the number of generations required for a random population to achieve a certain goal. The calculations were set up such that the studied system was evolving towards goals that change over time. They observed that changing the goals was in some cases greatly enhancing the speed of evolution with respect to evolutionary rates toward a "fixed" goal. Therefore, varying goals pushed the population away from the local fitness maxima, guiding them towards evolvable and modular solutions.

No experimental work on the role of the environment in fully synthetic systems has been reported yet. The role of environment on bacteria was studied experimentally by Lenski and coworkers in 1990.[4] An ancestral line of bacteria, maintained at 37 °C for 2000 generations, was used to prepare six lines divided in two sets, of which one is the high temperature experiment (42 °C) and the other is the control, kept at the same temperature. After maintaining the lines for 100, 200 and 400 generations at 42 °C, the bacteria were left competing at this temperature with the ancestral line. The authors observed that the fitness value increased by 7-8% after 200 generations with respect to the ancestral line. The control showed instead no significant

improvement in fitness. Therefore a change in environmental conditions resulted in a response from the bacteria towards an increase in fitness through mutations. Based on these observations, it is likely that the environment has also played an important role in the origin of life, pushing the first complex systems towards mutations that could have led to an increase in global fitness.

Previously, we observed that a change in the structure of the peptide building blocks synthesized in our group was leading to the emergence of different self-replicators.[5] We reasoned that external stimuli, like a change in the environment, may lead to the emergence of different self-replicators from the same building block. We decided to test the effects of a cosolvent in the system we developed and reported in chapter 1, in order to alter the interaction energy between the peptides and thereby change the size of the emerging self-replicator. Such finding would allow us to identify systems in which differently-sized replicators would compete for the same building block.

2.2. Results and Discussion

2.2.1. Identification of a suitable cosolvent to dictate the emergence of new species

To start our study on the effects of cosolvents on the emergence of new self-replicators, we focused on mixtures of water and 2,2,2-trifluoroethanol (TFE), as the latter cosolvent is known to enhance secondary structure formation in proteins and peptides.[6] Building blocks **1a-c** (Figure 2.1a,d,g) were used to test the cosolvent because of the different replicators size that emerges from dynamic molecular networks formed from these building blocks in water: in libraries composed of **1a**, the hexamer arises as a self-replicator, while libraries composed of either **1b** or **1c** give replicating octamers. The replicator size is influenced by the hydrophobicity of the peptide chain of the building blocks.[5]

Experiments were set up by allowing building blocks **1a-c** to oxidize in the presence of oxygen from the air in stirred borate buffer solutions (50 mM, pH 8.2) containing different percentages of TFE. We monitored the composition of the mixtures over time and observed that, for phenylalanine peptide **1a**, only trimer and tetramer macrocycles were formed (Figure 2.1b,c). However, dramatically different results were obtained for the less hydrophobic peptides **1b** and **1c**, depending on the percentage of TFE applied. Molecular networks made from alanine peptide **1b** gave rise to the predominant emergence of octamer self-replicator in the presence

of 10% v/v of cosolvent, similar to the behavior in pure water (Figure 2.1e). However, upon increasing the percentage of TFE from 10 to 13% (Figure 2.1f) hexameric macrocycles were the dominant species. Similar behavior was observed for libraries set up with the least hydrophobic building block **1c**, but at a somewhat larger percentage of TFE. Until 15% v/v of this cosolvent the octamer emerged as the predominant replicator (Figure 2.3h), like in pure water, while at 17% TFE the hexamer dominated (Figure 2.1i). The total peak area of the libraries was also monitored over time: we observed that the values are not affected by the composition of the library, confirming that we detect all library material and that the molar absorbivity of the building blocks are essentially independent of the macrocycle in which the building block resides. (Figure 2.2).

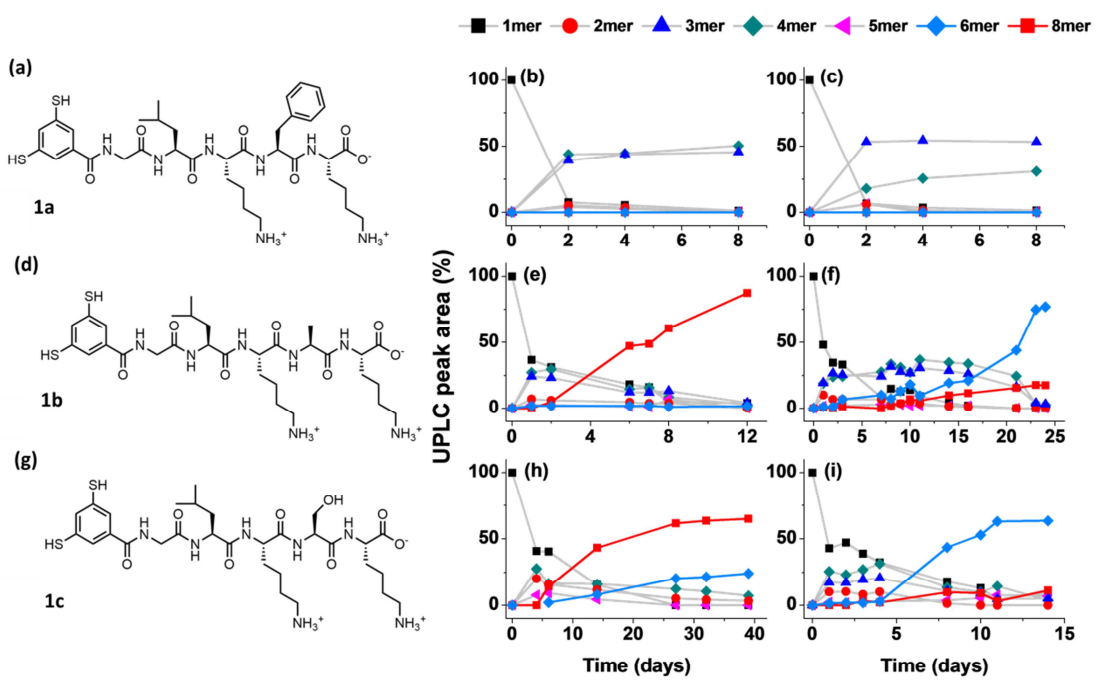


Figure 2.1. Structure of building block (a) **1a**; (d) **1b**; (g) **1c**. Change in product distributions of DCLs (3.8 mM dithiol building block in 50 mM borate buffer pH 8.2) made from (b) phenylalanine-containing peptide **1a** at 10% v/v TFE; (c) **1a** at 30% v/v TFE; (e) alanine-containing peptide **1b** at 10% v/v TFE; (f) **1b** at 13% TFE; (h) serine-containing peptide **1c** at 15% TFE; (i) **1c** at 17% TFE.

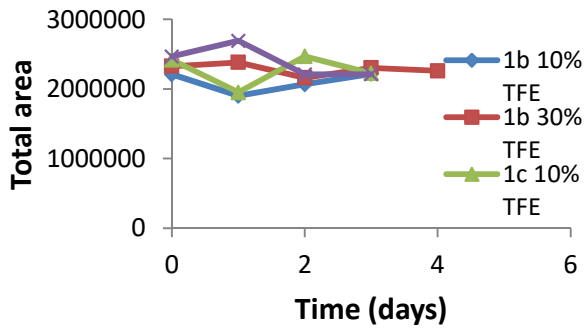


Figure 2.2. Total UPLC peak area of DCLs made from peptides **1b-c** monitored over time. The libraries were oxidized at day 0 at 70% with freshly prepared sodium perborate (38 mM), such that the dominant replicator (hexamer for 30% TFE solutions, octamer for 10% TFE solutions) could become the main species of the libraries already at day 3.

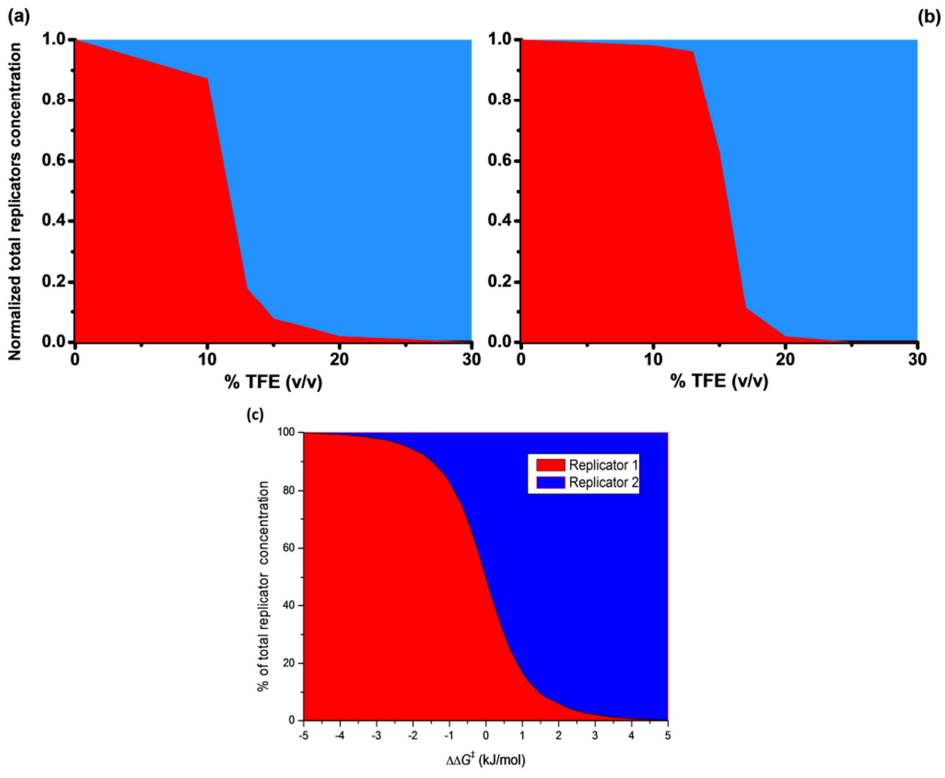


Figure 2.3. Summary of the experiments at different percentages of TFE (up to 30%) for (a) building block **1b** and (b) building block **1c**. The area in red indicates the relative peak area of the cyclic octamer, while the blue area shows the relative peak area of the cyclic hexamer. The octamer is the predominant replicator (a) until 10% TFE or (b) until 15% TFE. (c) Proportion of two competing replicators in a system where they compete for a common building block as a function of the difference in Gibbs activation energy of the replication reaction between the two replicators. The relationship between the proportion of replicators is a strongly non-linear function of the difference in Gibbs energy of activation. Equations can be found at page 52.

We performed similar experiments for a range of different TFE fractions. The final ratios between octamer and hexamer macrocycles of **1b** and **1c** as a function of the percentage TFE in the reaction medium are shown in Figure 2.3a,b, demonstrating that the switch from octamer to hexamer is a strongly non-linear function of the solvent composition. Such pronounced non-linear behavior is explained by the fact that the replicators are capable of exponential growth.[7] Simulations of a simplified model of two replicators competing for a common building block shows a similar sudden transition from one replicator to the other as the relative rate of replication changes (Figure 2.3c). It seems plausible that this transition is caused by the strengthening of peptide-peptide interactions resulting from the use of TFE as cosolvent. MD simulations by Mark et al.[8] suggest that preferential solvation of peptides by TFE enhances intrapeptide hydrogen bonding by reducing competition by water and providing a low dielectric environment. TFE shows only limited association with hydrophobic residues of the peptides, so hydrophobic interactions between the peptides are not significantly affected. With the TFE-mediated strengthening of the interactions between peptide building blocks, stacking becomes feasible for smaller macrocycles. To prove this hypothesis, libraries were set up in different water/ethanol mixtures of different compositions (Figure 2.4).

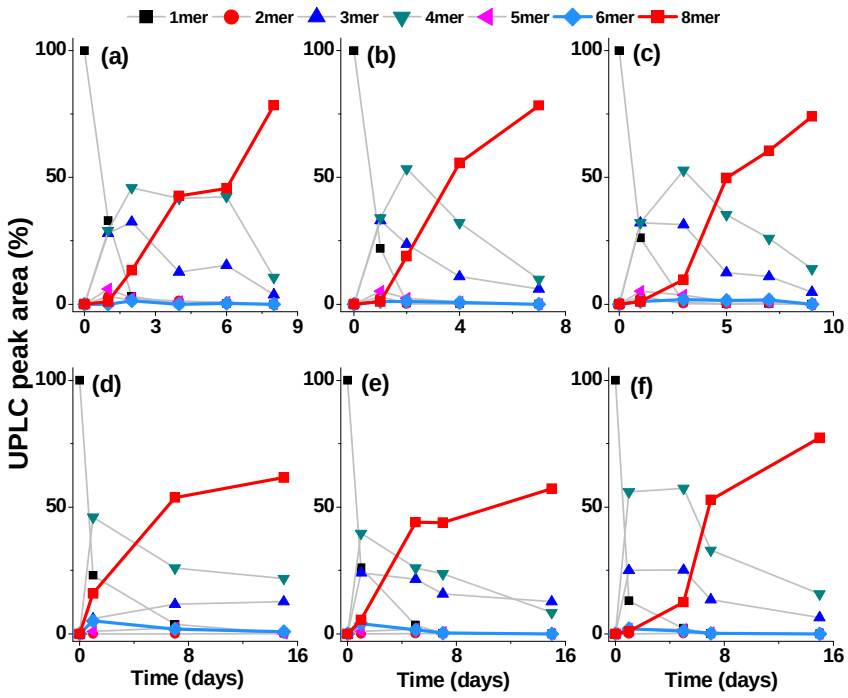


Figure 2.4. Evolution of the product distribution of DCLs made from 3.8 mM dithiol building block **1b** and **1c** in 50 mM borate buffer pH 8.2 in the presence of (a,d) 10%; (b,e) 15%; (c,f) 30% v/v ethanol. The predominant replicator is the octamer.

In all cases the octameric self-replicator emerged, but at a somewhat reduced rate compared with libraries set up in borate buffer only. No traces of hexameric self-replicator could be detected. Therefore the fluorine atoms are demonstrated of key importance in order to give rise to the hexamer. We also checked whether hexameric macrocycles could emerge from mixtures of water/hexafluoroisopropanol (HFIP) ranging from 5 to 10% v/v (Figure 2.5). However, in all the libraries we did not observe the emergence of any replicator, not even the octameric species which was emerging when using ethanol as cosolvent. This is probably due to the further stabilization of trimeric and tetrameric species by the larger number of fluorine atoms displaced by HFIP.

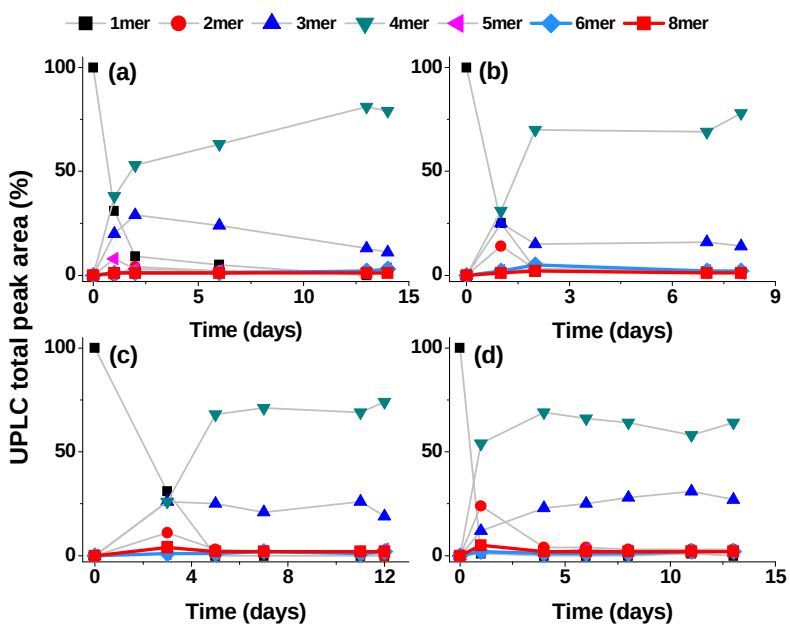


Figure 2.5. Evolution of the product distribution of DCLs made from 3.8 mM dithiol building block (a,b) **1b** and (c,d) **1c** in 50 mM borate buffer pH 8.2 in the presence of (a,c) 5%; (b,d) 10%; v/v hexafluoroisopropanol (HFIP). No traces of any hexameric or octameric replicators were observed.

2.2.2. Seeding experiments and characterization of the self-replicating species

We then verified whether the hexamers and octamers of **1b** and **1c** are self-replicators under different conditions. Samples were prepared by oxidizing **1b** or **1c** to produce solutions that were dominated by trimers and tetramers, therefore without the presence of any suspected replicators (Figure 2.6). Solutions were made in 10% TFE (for the octamers) and 30% TFE (for the hexamers) and to these solutions 10% of octamer or hexamer seeds were added. We then monitored the rate of growth of octamer or hexamer with UPLC and compared it with the corresponding rate in the absence of seed. The results are shown in Figure 2.7 and demonstrate that the addition of the suspected replicators indeed enhanced the rate of their own formation, confirming that both octamers and hexamers are self-replicators, i.e. capable of catalyzing the rate of their own formation.

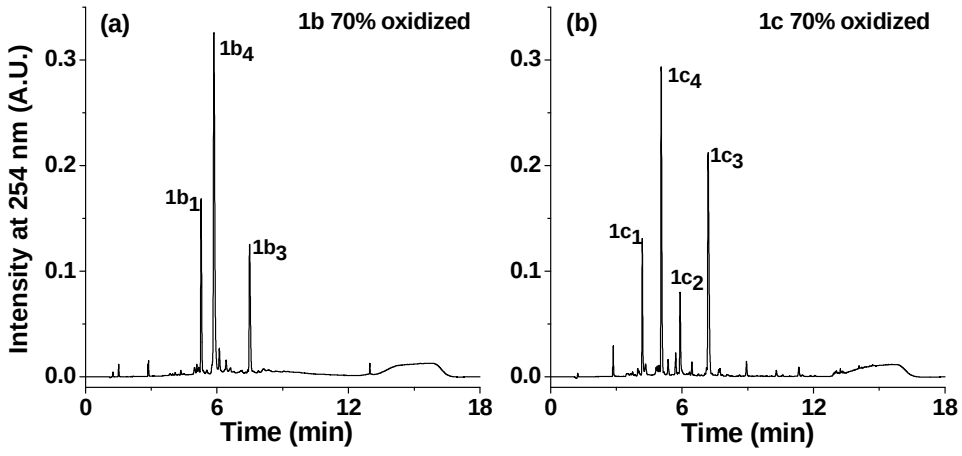


Figure 2.6. UPLC traces (monitored at 254 nm) of the product mixture obtained by 70% oxidation with sodium perborate of peptide (a) **1b**; (b) **1c**. After checking the oxidation level, the libraries were seeded and cross-seeded with octamer and hexamer replicators.

Replication is accompanied by fiber formation, as evident from transmission electron microscopy (TEM) micrographs (Figure 2.8a-d). The hexamer fibers formed from **1b** and **1c** in the presence of 30% TFE showed substantial lateral association, as was observed previously for the octamers of these building blocks.[5] The fibers of the hexamers of **1b** and **1c** were further characterized by circular dichroism (Figure 2.8e,f) and thioflavin T fluorescence experiments (Figure 2.8g,h). Both techniques confirmed the presence of β-sheets.

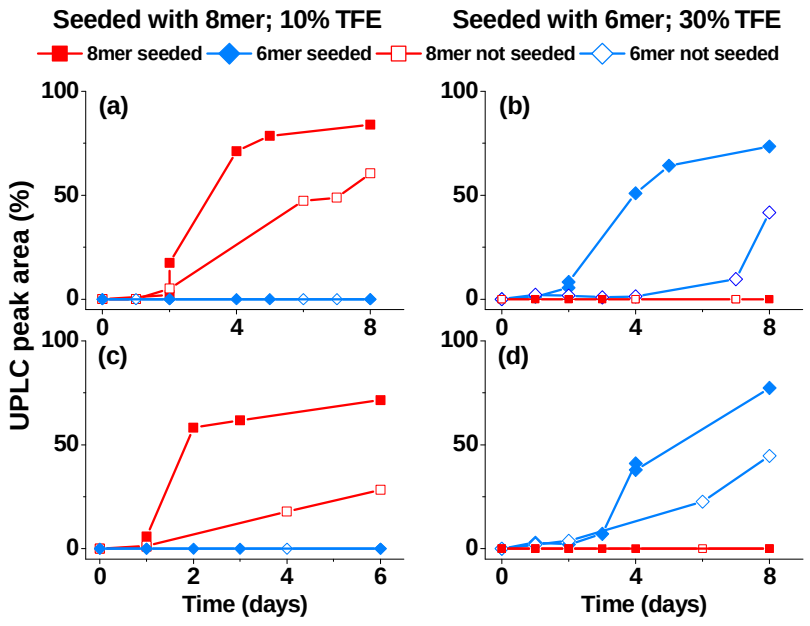
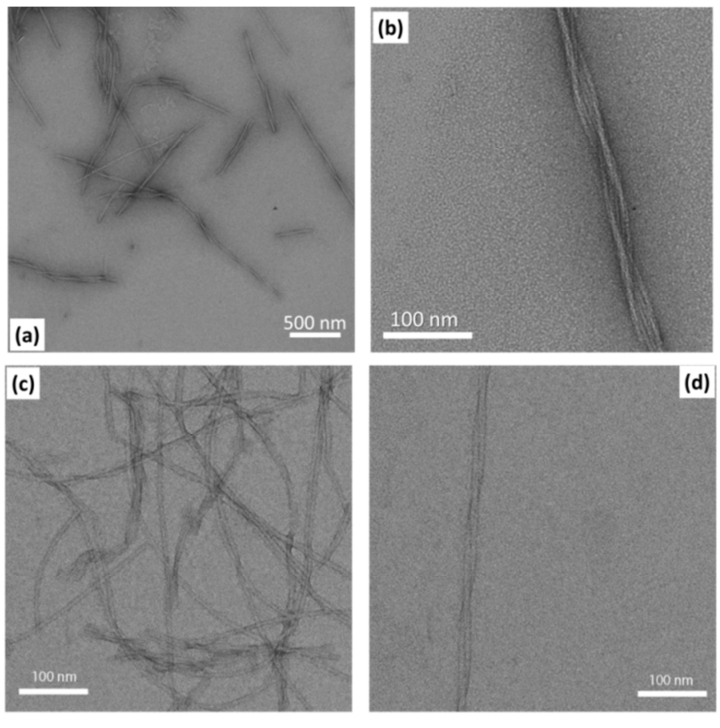


Figure 2.7. Seeding-induced growth of suspected self-replicating macrocycles under conditions that favor their formation. In all cases the libraries were seeded with 10% mol of suspected replicator. (a) Octamer seeded at day 2 in a 10% TFE solution containing **1b**; (b) Hexamer seeded at day 2 in a 30% TFE solution containing **1b**; (c) Octamer seeded at day 1 in a 10% TFE solution containing **1c**; (d) Hexamer seeded at day 4 in a 30% TFE solution containing **1c**. The data are compared with libraries that were not seeded (open symbols).[9] The peak area of the other species in the libraries is omitted for clarity.



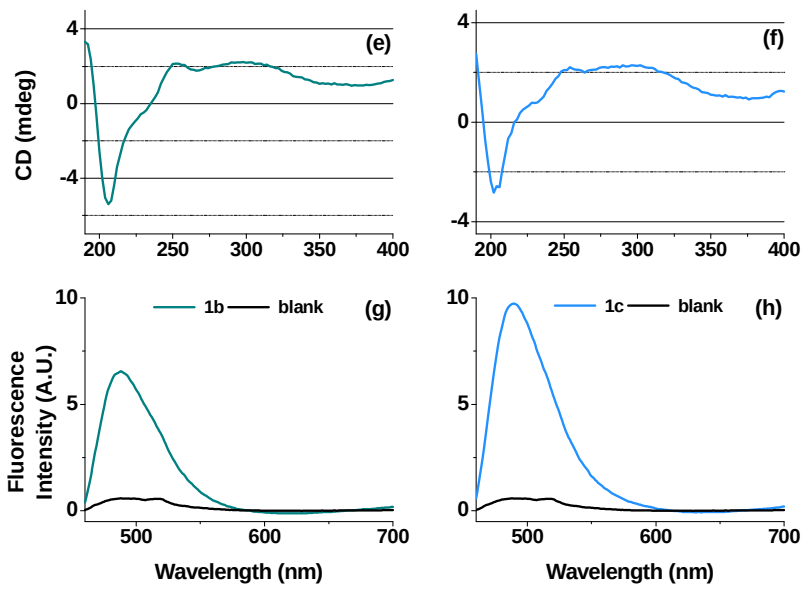


Figure 2.8. TEM micrographs taken from a DCL (borate buffer 50 mM, 30% v/v TFE, pH 8.2) made of peptide (a,b) **1b** and (c,d) **1c** that were dominated by hexamers. The images were obtained using negative staining (2% uranyl acetate). CD spectra of DCLs (borate buffer 50 mM, 30% v/v TFE, pH 8.2) made from peptides (e) **1b**; (f) **1c** and fluorescence emission spectra of solutions of ThT only (blank) and solutions of ThT made from peptide (g) **1b**; (h) **1c** that were dominated by hexameric replicators.

2.2.3. Cross seeding experiments

We also performed cross-seeding experiments to verify whether the hexamer and octamer macrocycles are able to replicate in environments in which they do not grow to significant concentrations in the absence of seeding. We seeded mixtures of trimer and tetramer macrocycles of **1b** or **1c** (prepared as described in the previous subchapter, Figure 2.6) with hexamer seed in 10% TFE and octamer seed in 30% TFE. In all the experiments, the seeded species was capable of self-replication, albeit to a very limited extent (Figure 2.9 and Figure 2.10). In all the experiments, the growth of the species that was added as a seed was much slower than the competing replicator, despite no seed of the latter had been added (Figure 2.9). The fact that we do not observe any consumption of the less favored replicators, also not during later stages of replication, suggests that the macrocycles that are part of fibers do not readily undergo disulfide exchange and appear to be kinetically trapped. Thus, even when the competing octamer replicator was added as a seed, the hexamer of **1b** and **1c** emerged as the predominant replicator in 30% TFE libraries (Figure 2.9b,d). Conversely, the octamer became the dominant replicator in 10% TFE experiments (Figure 2.9a,c) even when those libraries were seeded with competing hexameric replicators. Most surprisingly, the growth of the replicator `native` to a specific environment appears to benefit from seeding by its competing "non-

native" replicator, suggesting that cross-catalysis is taking place. The final compositions of the solutions of all eight seeding experiments on **1b** and **1c** are summarized in Figure 2.10 and compared with the compositions of the corresponding samples that were not seeded.

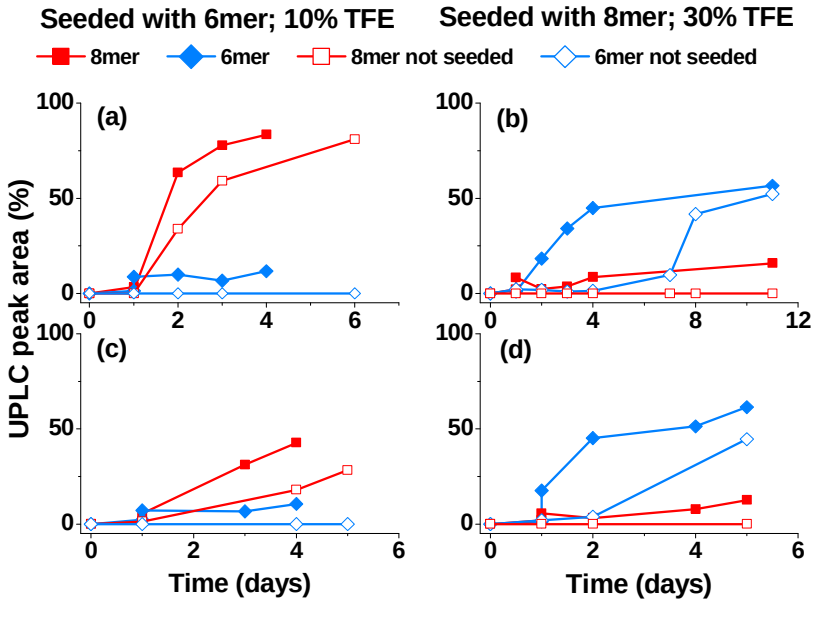


Figure 2.9. Seeding-induced growth of self-replicating macrocycles under conditions that *do not* favor their formation. (a) peptide **1b** at 10% TFE; (b) peptide **1b** at 30%TFE; (c) peptide **1c** at 10% TFE; (d) peptide **1c** at 30% TFE. The libraries were seeded on day 1 with 10 mol% of (a,c) hexamer; (b,d) octamer. The data are compared with libraries that were not seeded (open symbols).[9]

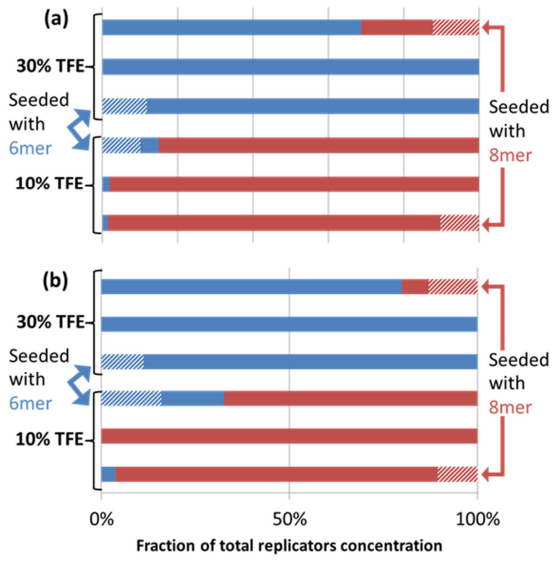


Figure 2.10. Summary of the seeding and cross-seeding experiments for (a) building block **1b** and (b) building block **1c**, compared with non-seeded experiments. The hexamer is shown in blue and the octamer in red. The striped areas correspond to the amount of added seed.

2.3. Conclusions

The obtained results capture for the first time at a molecular level the role of the environment in the competition between two replicating species for a common building block in a fully synthetic system. Perhaps not surprisingly, the environment plays a decisive role in the emergence of self-replicators from complex mixtures. The proportion of two competing replicators was found to be a strongly non-linear function of the solvent composition, since replicators are not able to emerge in a less favorable environment. With TFE-mediated strengthening of the interactions between peptide building blocks stacking becomes feasible for smaller macrocycles: in the presence of TFE, oxidation of building block **1a** (which formed hexamers in water) resulted in the formation of trimers and tetramers only. Instead, building blocks **1b** and **1c**, that give self-replicating octamers in water, stack in hexamers with appropriate amounts of TFE. The concentration of TFE required for the changeover was higher for the less hydrophobic and therefore less strongly interacting serine-containing building block **1c** as compared to alanine-containing **1b**. These trends are summarized in Figure 2.11, which shows qualitatively the strength of the peptide/peptide interactions for **1a-c** in different water/TFE mixtures. The horizontal lines separate the areas in which specific macrocycles dominate.

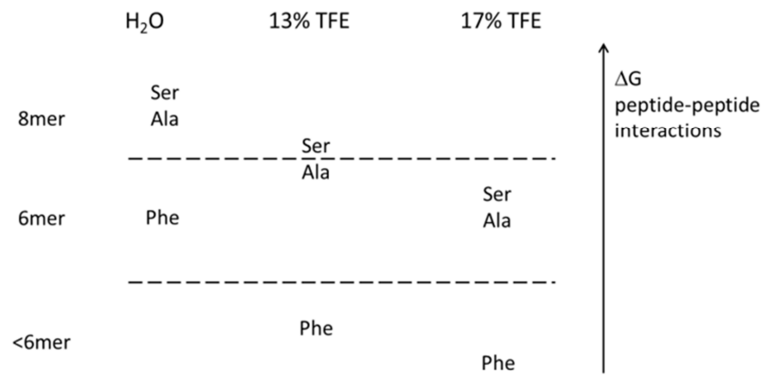


Figure 2.11. Estimated strength of the peptide-peptide interactions for the different peptides in water/TFE mixtures. Introduction of TFE strengthens the peptide-peptide interactions, making stacking feasible for smaller macrocycles.

In a more general context, this finding opens new possibilities in systems chemistry,[10] specifically for investigating Darwinian evolution at the molecular level, since the results show the fitness of competing, fully synthetic replicators can be tuned by changing the environment. It is also possible to define niches by creating specific environments and develop out-of-

equilibrium systems (as the one that will be described in Chapter 3) where the replicators can compete for the same resources.

2.4. Acknowledgments

Geert R. Hoogeboom is acknowledged for the experiments carried out in ethanol and hexafluoroisopropanol. Prof. Sijbren Otto is acknowledged for the kinetic model of two replicators using Berkeley-Madonna software package.

2.5. Experimental Section

2.5.1. Materials

Water was doubly distilled prior to use. Boric acid and potassium hydroxide utilized for the preparation of buffers and pH adjustment were obtained from Acros Organics and Merck Chemicals, respectively. 2,2,2-Trifluoroethanol was purchased from Sigma Aldrich. Ethanol was purchased from J. T. Baker. Acetonitrile (UPLC grade), water (UPLC grade) and trifluoroacetic acid were purchased from Biosolve BV.

2.5.2. General Methods

Peptide synthesis

Peptides **1a-c** were synthesized by Cambridge Peptides Ltd (Birmingham, UK) from 3,5-bis(tritylthio)benzoic acid, which was prepared via a previously reported procedure.[11] All peptides showed purity higher than 85%. Impurities were mostly due to oxidation of thiols to disulfides (i.e. dimer, trimer).

Library preparation and sampling

Building blocks **1a-c** were dissolved to a concentration of 3.8 mM in borate buffer (50 mM, pH 8.2) and 2,2,2-trifluoroethanol in 10, 13, 15, 17, 20, 25 and 30% v/v (10% and 30% only for **1a**). Where necessary the pH of the solution was adjusted by the addition of 2.0 M KOH solution such that the final pH was 8.2. The volume of each library was 400 µL. Each solution was allowed to equilibrate in an HPLC vial (12 x 32 mm) with a Teflon-lined snap cap. All the samples contained a cylindrical micro-stirrer bar (2 x 5 mm, Teflon-coated, purchased from VWR) and were stirred at 1200 rpm using an IKA RCT basic hotplate stirrer. All the experiments were performed at ambient temperature. A small aliquot of each sample was

removed to another vial and diluted 20 times with double distilled water prior to UPLC or LC-MS analysis.

Seeding experiments

Two libraries containing building blocks **1b-c** in 10% and 30% TFE were prepared accordingly to the procedure described above. The solutions were oxidized to 70% with sodium perborate (freshly prepared, 38 mM) such that they contained mostly monomer, trimer and tetramer. The libraries were then split in two parts. To one part prepared with 10 % TFE, a small amount (10 mol %) of pre-existing octamer was added. The same amount of hexamer was instead added to one of the two libraries prepared with 30% TFE. All four libraries were monitored by UPLC. The preparation of the sample is described above.

Cross seeding experiments

For peptides **1b-c** two solutions containing mostly monomer, trimer and tetramer at 10% and 30% TFE were prepared according to the procedure described in the seeding experiments paragraph. The libraries were then split in two parts. To one of the two libraries prepared with 10 % TFE a small amount (10 mol %) of pre-formed hexamer was added. The same amount of octamer was instead added to one of the two libraries prepared with 30% TFE. All four libraries were monitored by UPLC.

Negative staining Transmission Electron Microscopy

A small drop (5 µL) of sample was deposited on a 400 mesh copper grid covered with a thin carbon film (Agar Scientific). After 30 seconds, the droplet was blotted on filter paper. The sample was then stained twice (4 µL each time) with a solution of 2% uranyl acetate deposited on the grid and blotted on the filter paper after 30 seconds each time. The grids were observed in a Philips CM120 cryo-electron microscope operating at 120 kV. Images were recorded on a slow scan CCD camera.

Thioflavine T (ThT) fluorescence

Sample aliquots were diluted to a concentration of 76 µM with respect to peptides **1b-c** with additional potassium borate buffer (50 mM, pH 8.2). The diluted sample (3.2 µL) was added to a ThT solution (22 µM, 26.8 µL) in potassium borate buffer (50 mM, pH 8.2) and incubated for 5 minutes. The solution was diluted with additional potassium borate buffer (50 mM, pH 8.2, 100 µL) and transferred into HELMA 10 x 2 mm quartz cuvettes. The fluorescence was

measured on a JASCO FP6200 spectrophotometer by excitation at 440 nm (5 nm slit width) and emission from 460 to 700 nm (5 nm slit width, 3 repeats averaged).

Circular dichroism (CD)

Spectra were obtained at 20 °C using a JASCO J715 spectrophotometer (range 190-400 nm, pitch: 2 nm, bandwidth: 5 nm, response: 2 s, speed: 50 nm/min, continuous scanning) and HELMA 10 x 2 mm quartz cuvettes. All reported spectra are averages of 3 repeats. Solvent spectra were subtracted from all spectra. All spectra were obtained using samples diluted to 8 μM (with respect to building block concentration).

UPLC and LC-MS analyses

UPLC analyses were performed on a Waters Acquity H-class equipped with diode array UV/Vis detector. LC-MS analyses were performed on a Xevo G2 UPLC/TOF with ESI ionization, manufactured by Waters. All analyses were performed at 35 °C using a reversed-phase UPLC column (Phenomenex Aeris Peptide, 2.1 x 150 mm; 1.7 μm). UV absorbance was monitored at 254 nm. Positive-ion mass spectra were acquired using electro-spray ionization. Injection volume 5 μL of freshly aliquoted sample; column temperature 35 °C; flow rate 0.3 mL/min.

UPLC methods

Solutions containing peptides **1a-c** and their oxidation products were analyzed using the following methods (linear gradient):

Solvent A: ULC/MS grade water purchased from Biosolve (0.1% trifluoroacetic acid added)

Solvent B: ULC/MS grade acetonitrile purchased from Biosolve (0.1% trifluoroacetic acid added)

Peptide **1b**

(Phenomenex Aeris Peptide column)

Time (min)	A%	B%
0	90	10
1	75	25
11	55	45
12	5	95
14	5	95
15	90	10
18	90	10

Peptide **1c**

(Phenomenex Aeris Peptide column)

Time (min)	A%	B%
0	90	10
1	75	25
11	65	35
12	5	95
14	5	95
15	90	10
18	90	10

LC-MS analysis of 1b

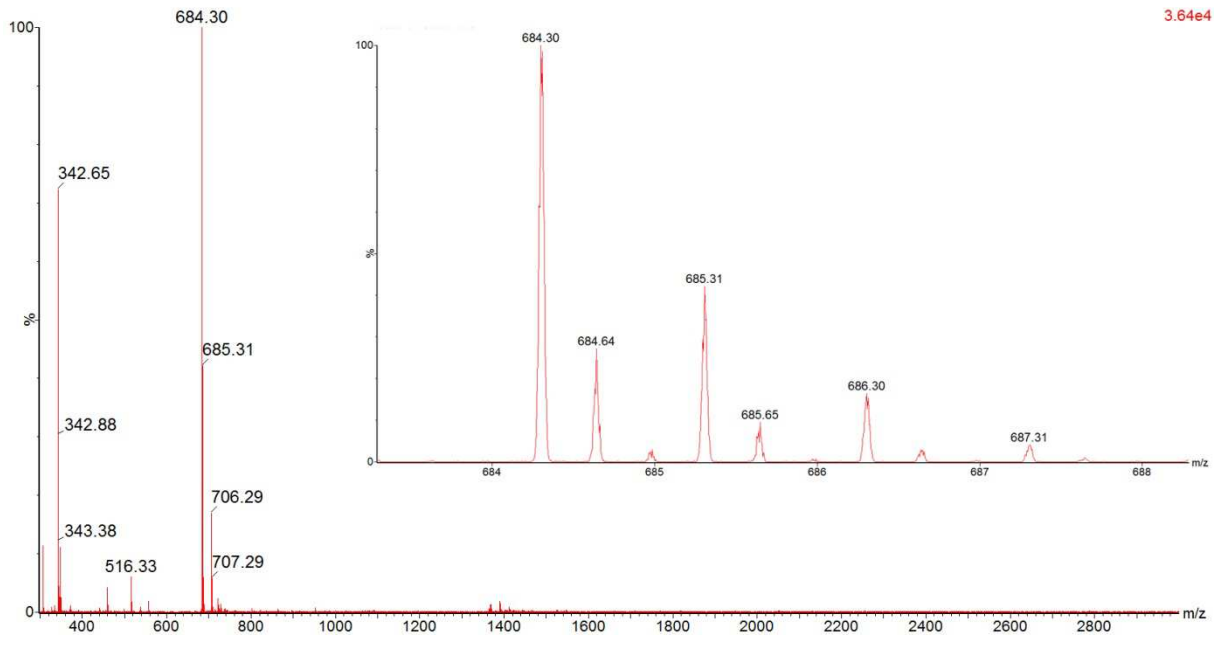


Figure 2.12. Mass spectrum of building block **1b** from the LC-MS analysis of a stirred library made from peptide **1b**. Calculated isotopic profile (species, abundance): 683.30 (M, 100.0%), 684.30 (M+1, 37.4%), 685.29 (M+2, 17.3%); m/z calculated: 684.31 $(M+1H)^+$; m/z observed: 684.30 $(M+1H)^+$. Inset: isotopic profile at m/z 684.30.

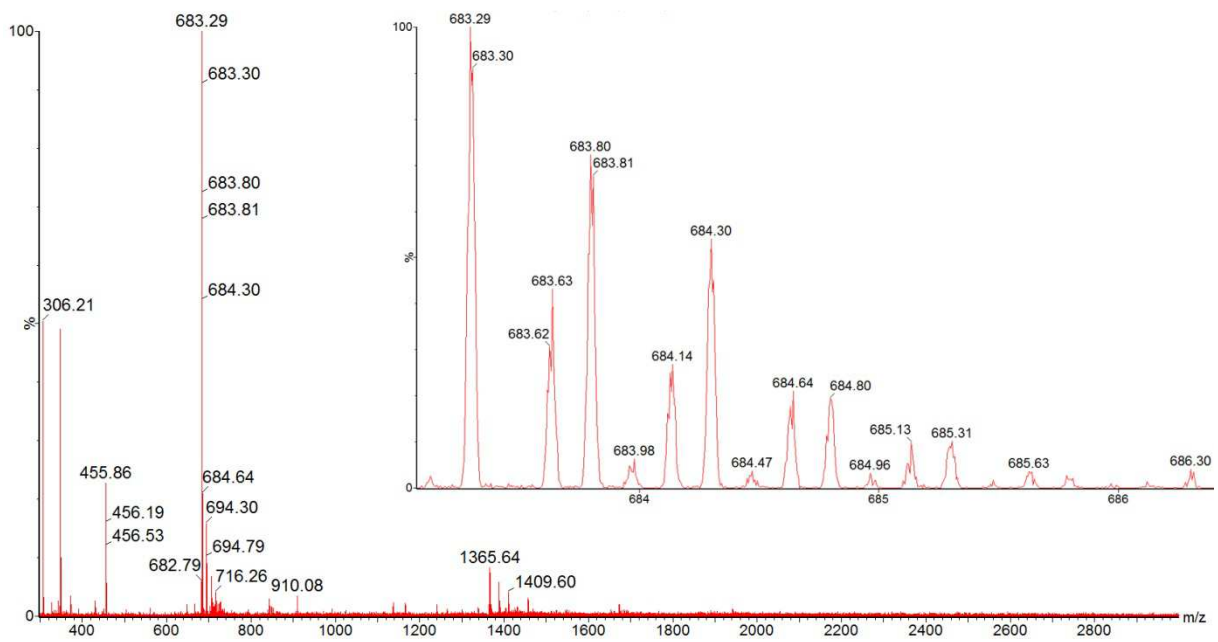


Figure 2.13. Mass spectrum of linear dimer $\mathbf{1b}_2$ from the LC-MS analysis of a stirred library made from peptide **1b**. Calculated isotopic profile (species, abundance): 1364.60 (M, 100.0%), 1365.60 (M+1, 74.9%), 1366.60 (M+2, 48.6%); m/z calculated: 683.31 $(M+2H)^{2+}$; m/z observed: 683.29 $(M+2H)^{2+}$. Inset: isotopic profile at m/z 683.29.

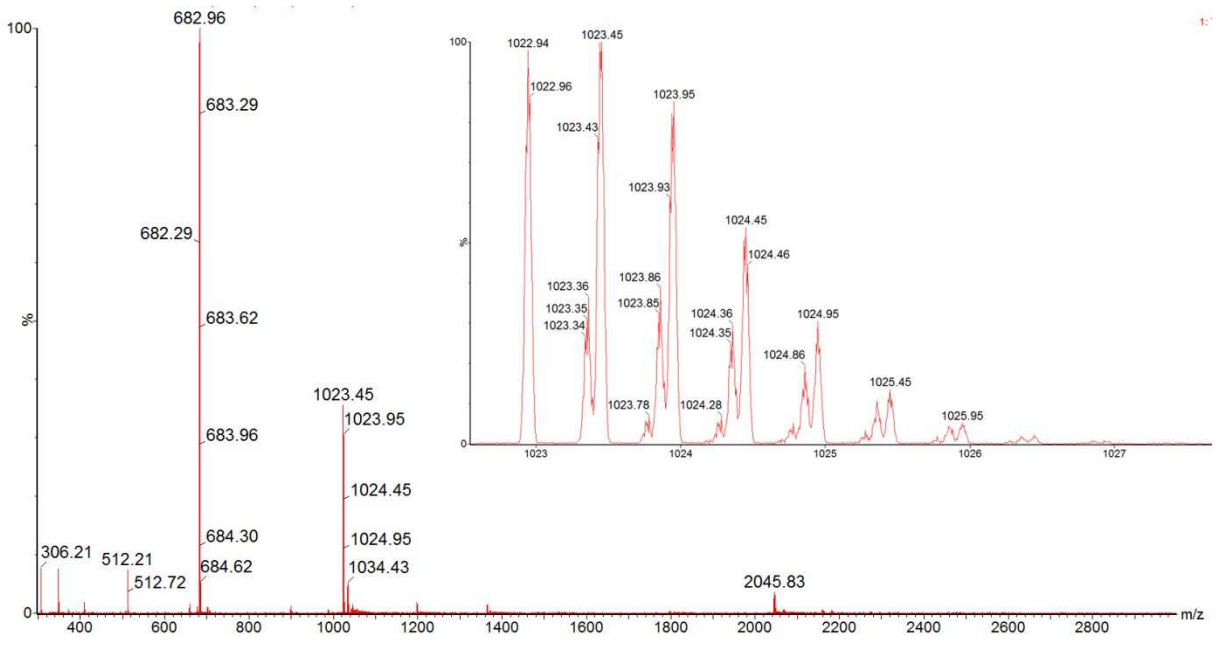


Figure 2.14. Mass spectrum of cyclic trimer **1b**$_3$ from the LC-MS analysis of a stirred library made from peptide **1b**. Calculated isotopic profile (species, abundance): 2044.90 (M+1, 100.0%), 2043.89 (M, 89.0%), 2045.90 (M+2, 83.6%); m/z calculated: 1023.45 $[(M+1)+2H]^{2+}$, 682.63 $[(M+1)+3H]^{3+}$; m/z observed: 1023.45 $[(M+1)+2H]^{2+}$, 682.96 $[(M+1)+3H]^{3+}$. Inset: isotopic profile at m/z 1023.45.

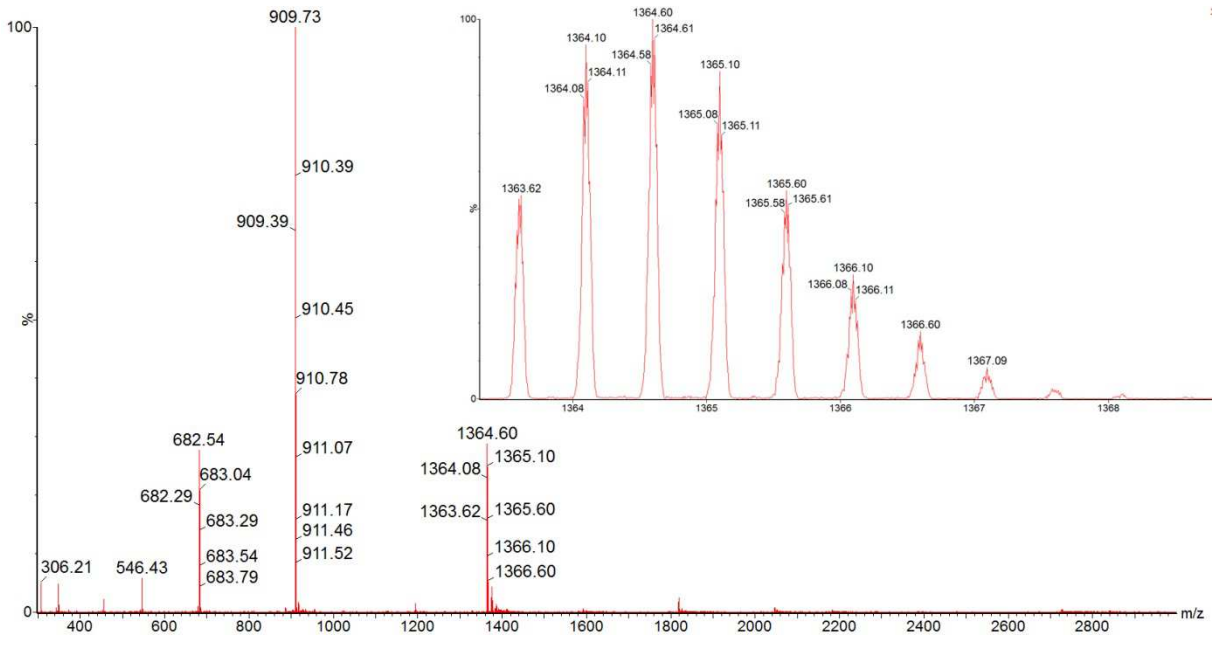


Figure 2.15. Mass spectrum of cyclic tetramer **1b**$_4$ from the LC-MS analysis of a stirred library made from peptide **1b**. Calculated isotopic profile (species, abundance): 2726.19 (M+1, 100.0%), 2725.19 (M, 77.0%), 2727.20 (M+2, 64.4%); m/z calculated: 1365.26 $[(M+1)+2H]^{2+}$, 909.73 $[(M+1)+3H]^{3+}$; m/z observed: 1364.60 $[(M+1)+2H]^{2+}$, 909.73 $[(M+1)+3H]^{3+}$. Inset: isotopic profile at m/z 1364.60.

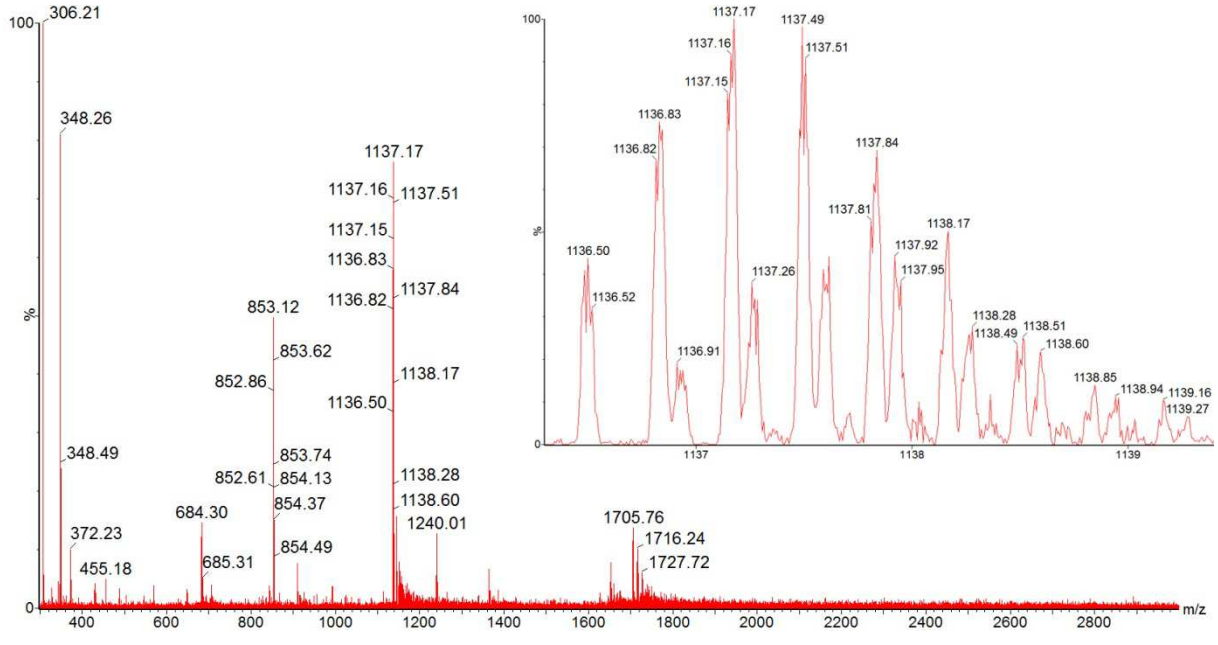


Figure 2.16. Mass spectrum of cyclic pentamer **1b**$_5$ from the LC-MS analysis of a stirred library made from peptide **1b**. Calculated isotopic profile (species, abundance): 3407.49 (M+1, 100%), 3406.49 (M, 61.6%), 3408.5 (M+2, 54.2%); m/z calculated: 1137.80 $[(M+1)+3H]^{3+}$, 853.85 $[(M+1)+4H]^{4+}$; m/z observed: 1137.17 $[(M+1)+3H]^{3+}$, 853.12 $[(M+1)+4H]^{4+}$. Inset: isotopic profile at m/z 1137.17.

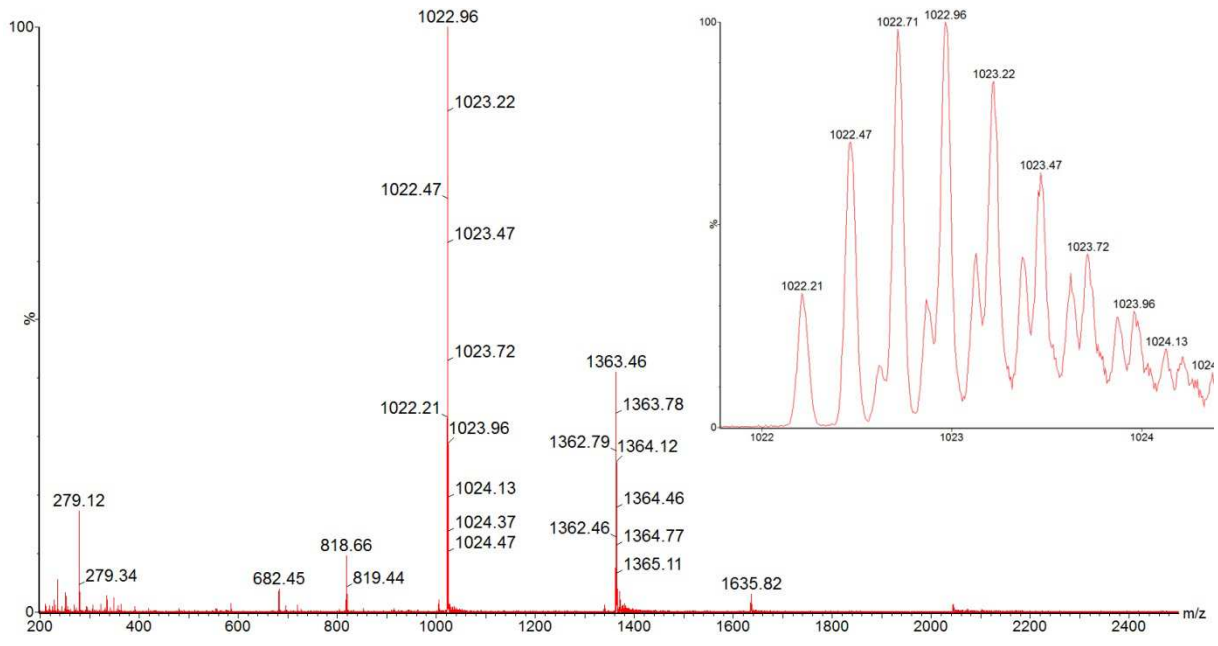


Figure 2.17. Mass spectrum of cyclic hexamer **1b**$_6$ from the LC-MS analysis of a stirred library made from peptide **1b**. Calculated isotopic profile (species, abundance): 4088.79 (M+1, 100%), 4089.79 (M+2, 96.8%), 4090.80 (M+3, 54.2%); m/z calculated: 1363.93 $[(M+1)+3H]^{3+}$, 1023.82 $[(M+1)+4H]^{4+}$; m/z observed: 1363.46 $[(M+1)+3H]^{3+}$, 1022.96 $[(M+1)+4H]^{4+}$. Inset: isotopic profile at m/z 1022.96.

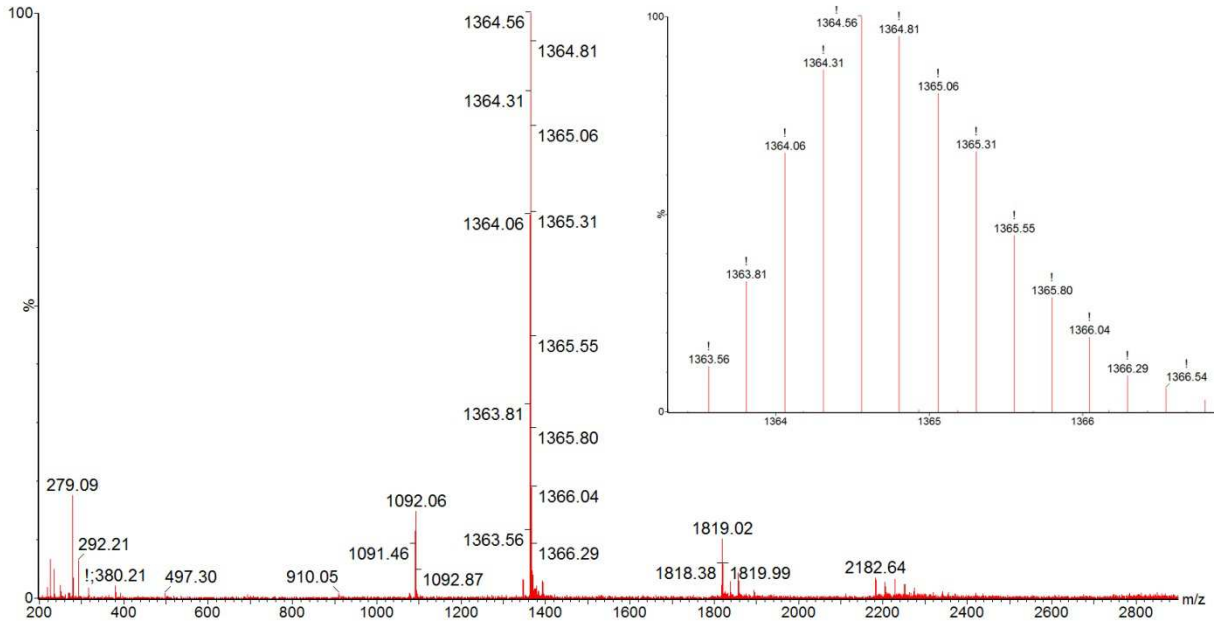


Figure 2.18. Mass spectrum of cyclic octamer **1b**$_8$ from the LC-MS analysis of a stirred library made from peptide **1b**. Calculated isotopic profile (species, abundance): 5452.39 (M+2, 100.0%), 5453.39 (M+3, 85.8%), 5451.40 (M+1, 77.4%); m/z calculated: 1364.1 $[(M+2)+4H]^{4+}$, 1092.41 $([M+2)+5H]^{5+}$; m/z observed: 1364.56 $[(M+2)+4H]^{4+}$, 1092.06 $[(M+2)+5H]^{5+}$. Inset: isotopic profile at m/z 1364.56.

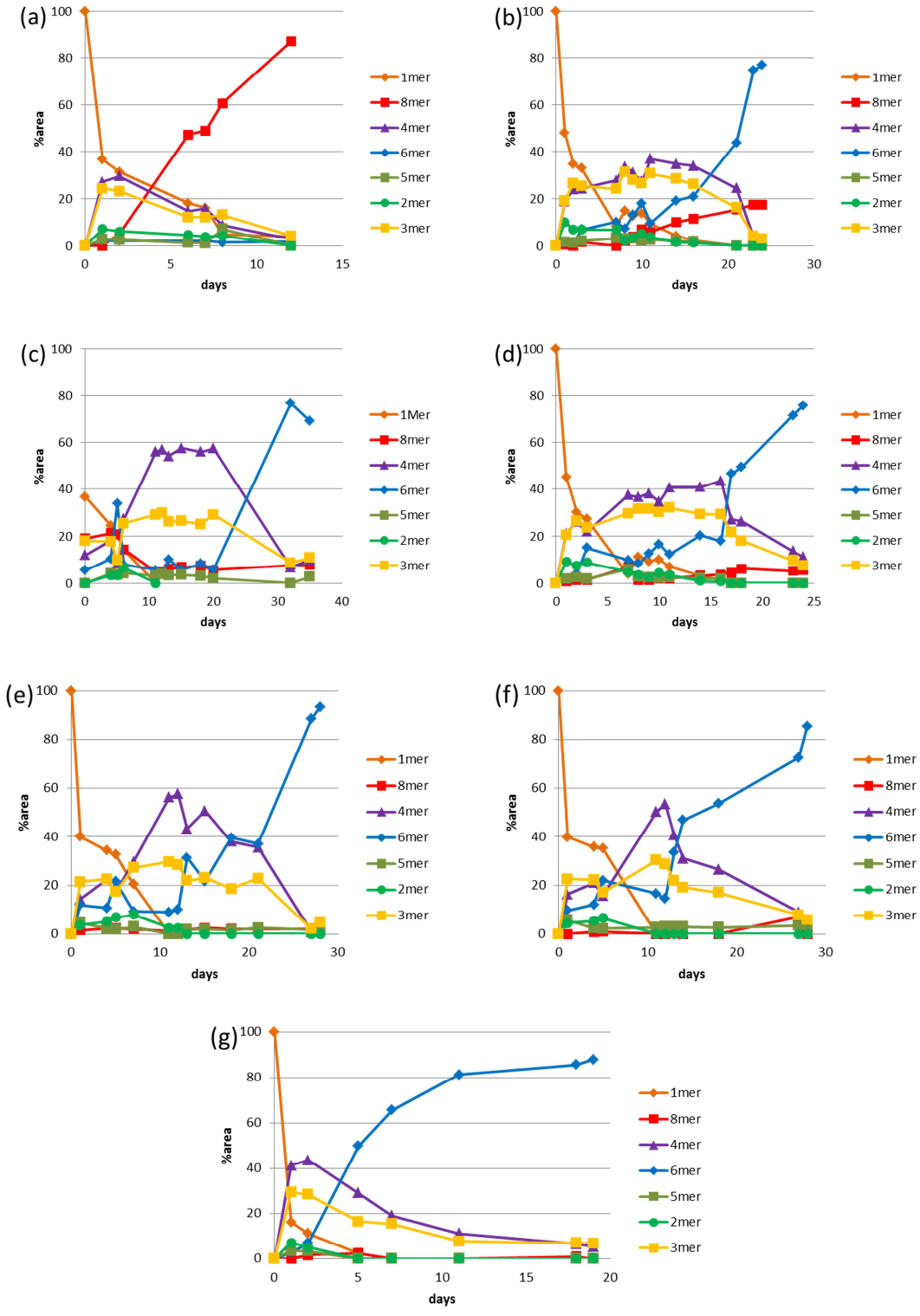


Figure 2.19. Evolution of the product distribution of DCLs (3.8 mM dithiol building block **1b** in 50 mM borate buffer pH 8.2) at different concentration of TFE. (a) 10%: the predominant replicator is the octamer; (b) 13%; (c) 15%; (d) 17%; (e) 20%; (f) 25%; (g) 30%. In graphs (b-g) the predominant replicator is the hexamer.

LC-MS analysis of 1c

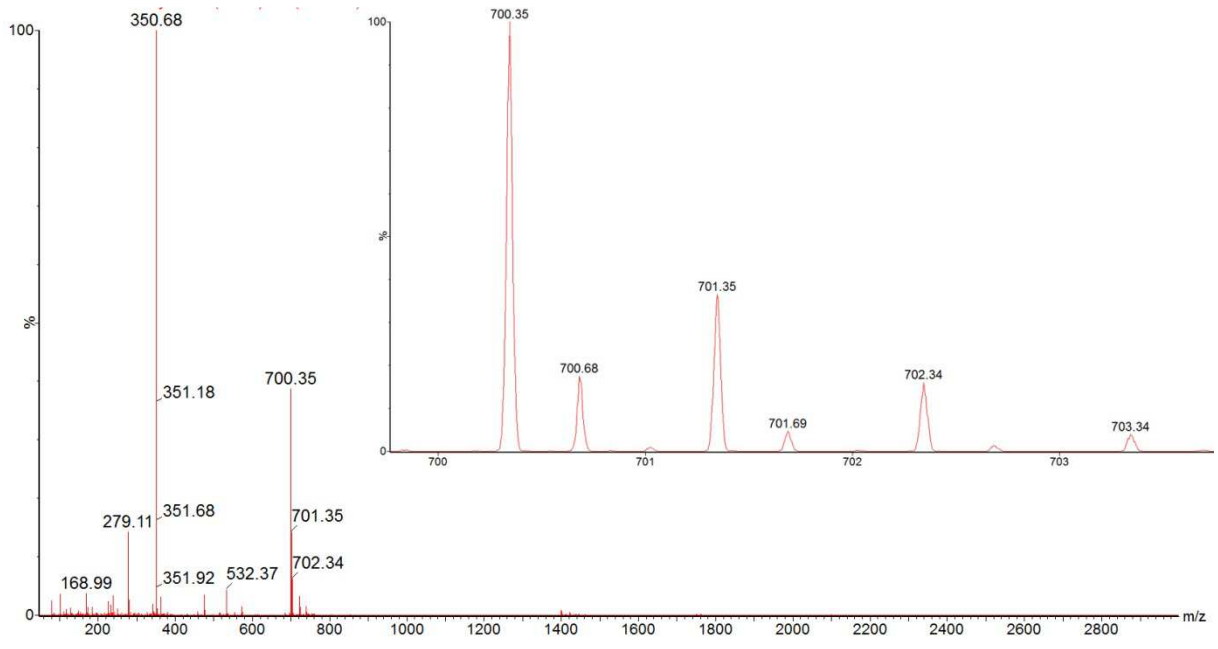


Figure 2.20. Mass spectrum of building block **1c** from the LC-MS analysis of a stirred library made from peptide **1c**. Calculated isotopic profile (species, abundance): 700.32 (M, 100.0%), 731.32 (M+1, 32.4%), 732.31 (M+2, 9%); m/z calculated: 700.32 $(M+H)^+$, 351.16 $(M+2H)^{2+}$; m/z observed: 700.35 $(M+H)^+$, 350.68 $(M+2H)^{2+}$. Inset: isotopic profile at m/z 700.35.

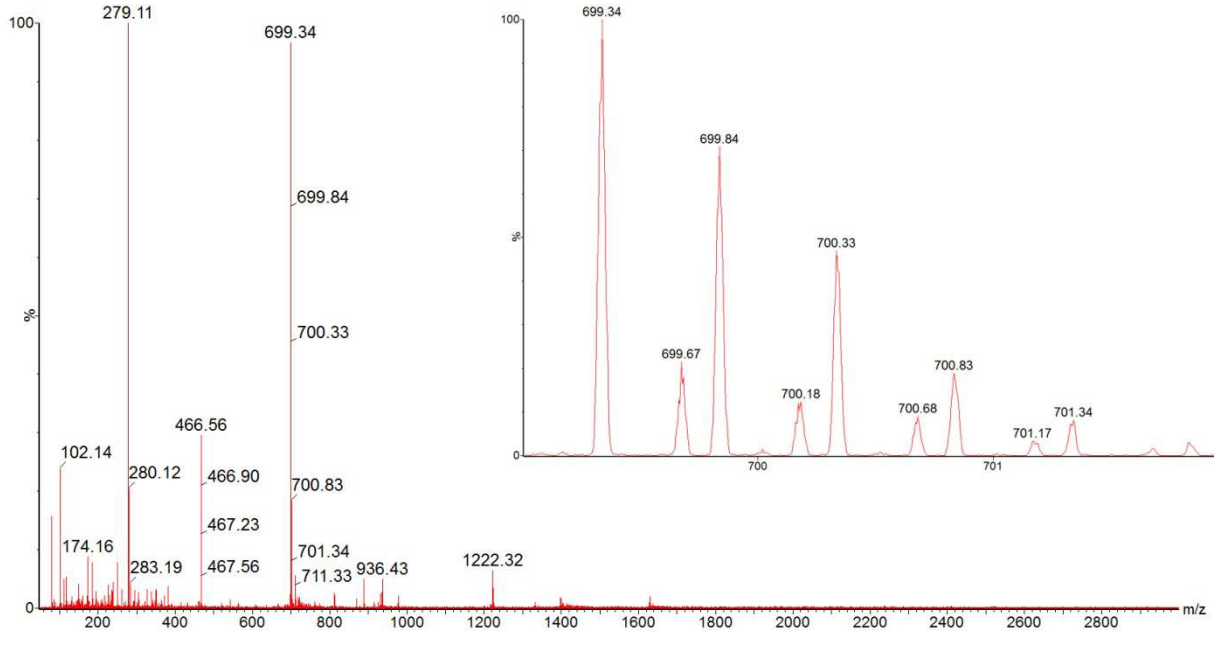


Figure 2.21. Mass spectrum of linear dimer $\mathbf{1c}_2$ from the LC-MS analysis of a stirred library made from peptide **1c**. Calculated isotopic profile (species, abundance): 1396.62 (M, 100.0%), 1397.62 (M+1, 75.0%), 1398.61 (M+2, 49.1%); m/z calculated: 699.31 $(M+2H)^{2+}$; m/z observed: 699.34 $(M+2H)^{2+}$. Inset: isotopic profile at m/z 699.34.

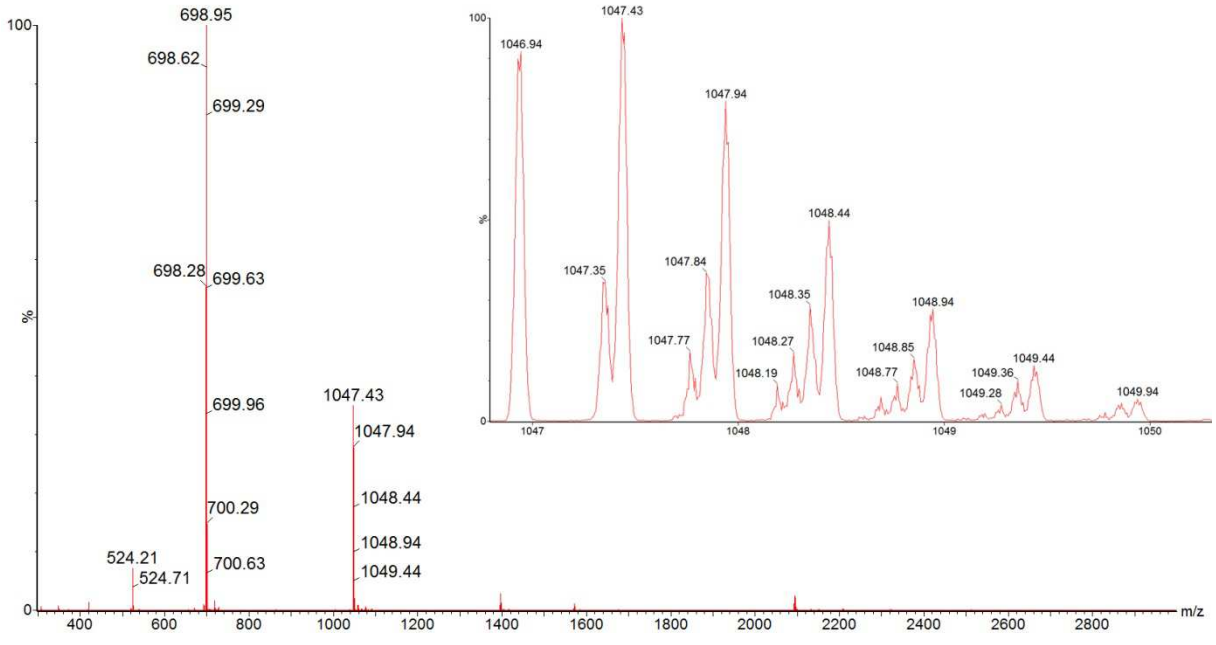


Figure 2.22. Mass spectrum of cyclic trimer **1c**$_3$ from the LC-MS analysis of a stirred library made from peptide **1c**. Calculated isotopic profile (species, abundance): 2092.96 (M+1, 100.0%), 2091.96 (M, 88.9%), 2093.96 (M+2, 84.2%); m/z calculated: 1047.46 $[(M+1)+2H]^{2+}$, 698.64 $[(M+1)+3H]^{3+}$, 524.24 $[(M+1)+4H]^{4+}$. m/z observed: 1047.43 $[(M+1)+2H]^{2+}$, 698.95 $[(M+1)+3H]^{3+}$, 524.21 $[(M+1)+4H]^{4+}$. Inset: isotopic profile at m/z 1047.43.

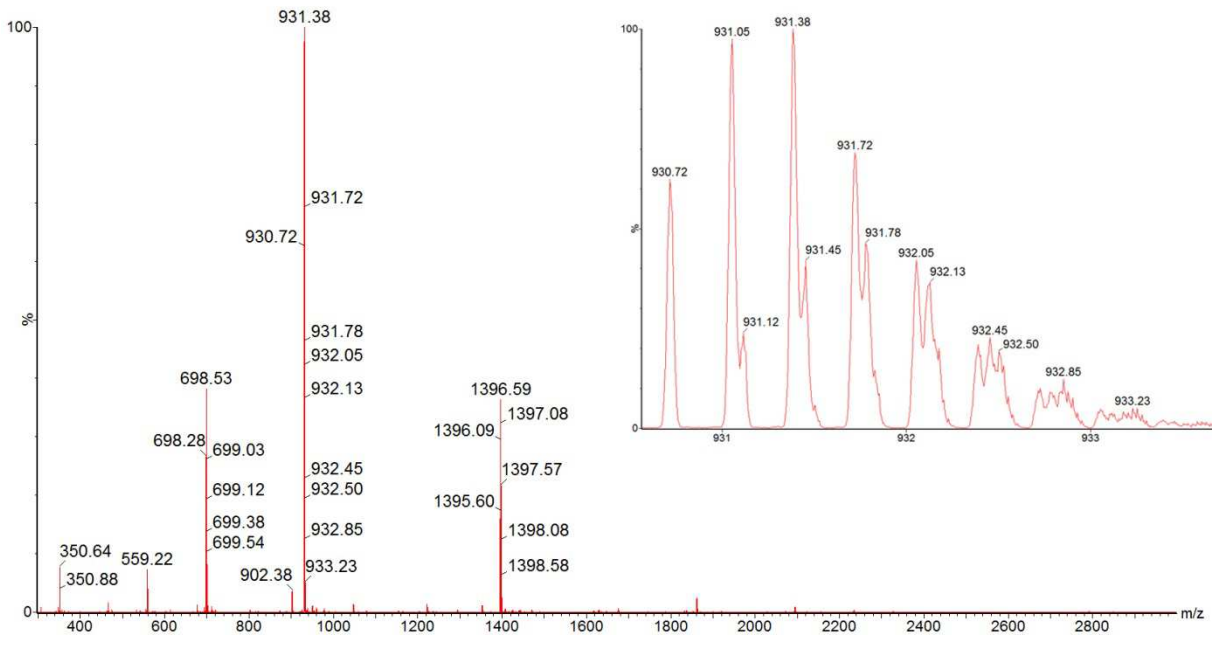


Figure 2.23. Mass spectrum of cyclic tetramer **1c**$_4$ from the LC-MS analysis of a stirred library made from peptide **1c**. Calculated isotopic profile (species, abundance): 2791.24 (M+2, 100.0%), 2790.24 (M+1, 97.1%), 2792.23 (M+3, 76.9%), 2789.23 (M, 64.8%); m/z calculated: 1396.12 $[(M+1)+2H]^{2+}$, 931.08 $[(M+1)+3H]^{3+}$, 698.55 $[(M+1)+4H]^{4+}$; m/z observed: 1396.59 $[(M+1)+2H]^{2+}$, 931.38 $[(M+1)+3H]^{3+}$, 698.53 $[(M+1)+4H]^{4+}$. Inset: isotopic profile at m/z 931.38.

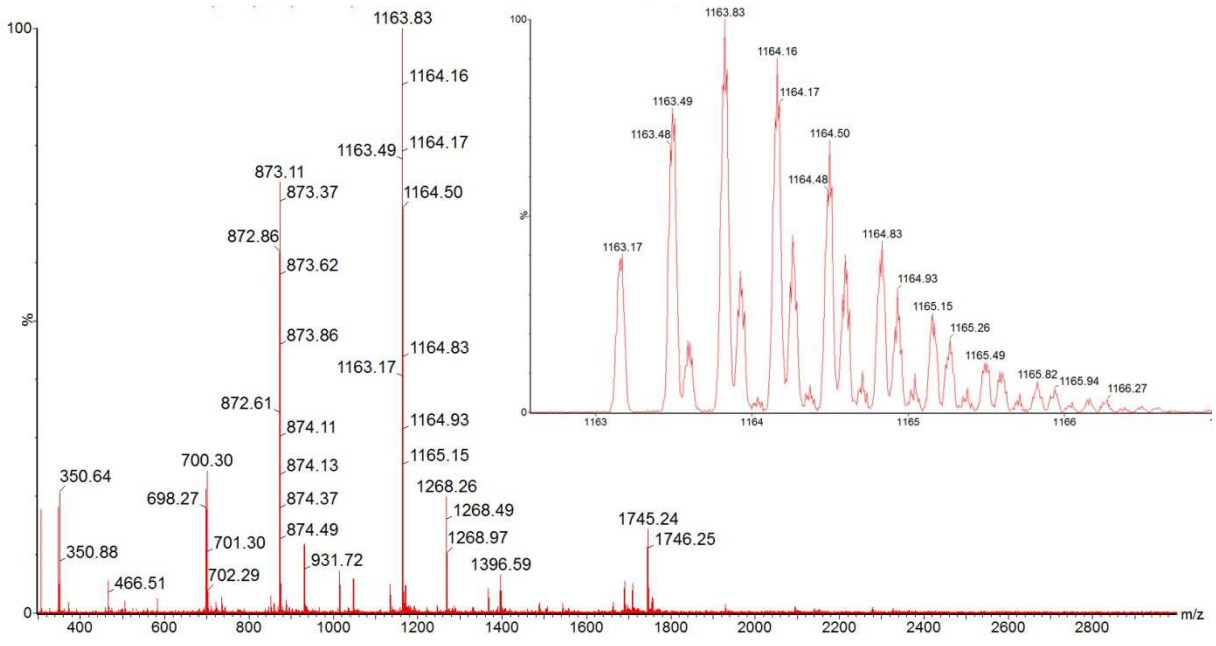


Figure 2.24. Mass spectrum of cyclic pentamer **1c**$_5$ from the LC-MS analysis of a stirred library made from peptide **1c**. Calculated isotopic profile (species, abundance): 3488.55 (M+2, 100.0%), 3489.55 (M+3, 91.0%), 3487.55 (M+1, 82.2%), 3490.54 (M+4, 68.2%), 3486.55 (M, 43.9%); m/z calculated: 1163.51 $[(M+1)+3H]^{3+}$, 872.88 $[(M+1)+4H]^{4+}$; m/z observed: 1163.83 $[(M+1)+3H]^{3+}$, 873.11 $[(M+1)+4H]^{4+}$. Inset: isotopic profile at m/z 1163.83.

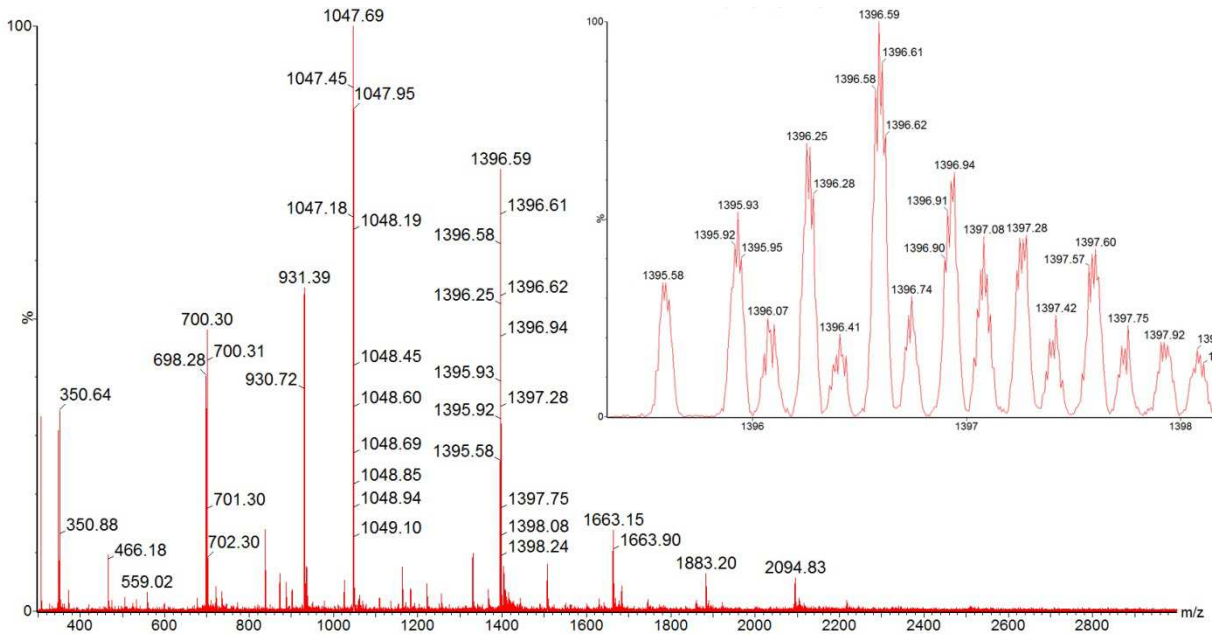


Figure 2.25. Mass spectrum of cyclic hexamer **1c**$_6$ from the LC-MS analysis of a stirred library made from peptide **1c**. Calculated isotopic profile (species, abundance): 4184.76 (M+1, 100.0%), 4185.76 (M+2, 96.8%), 4187.76 (M+4, 52.5%); m/z calculated: 1396.25 $[(M+2)+3H]^{3+}$, 1047.44 $[(M+2)+4H]^{4+}$; m/z observed: 1396.59 $[(M+2)+3H]^{3+}$, 1047.69 $[(M+2)+4H]^{4+}$. Inset: isotopic profile at m/z 1396.59.

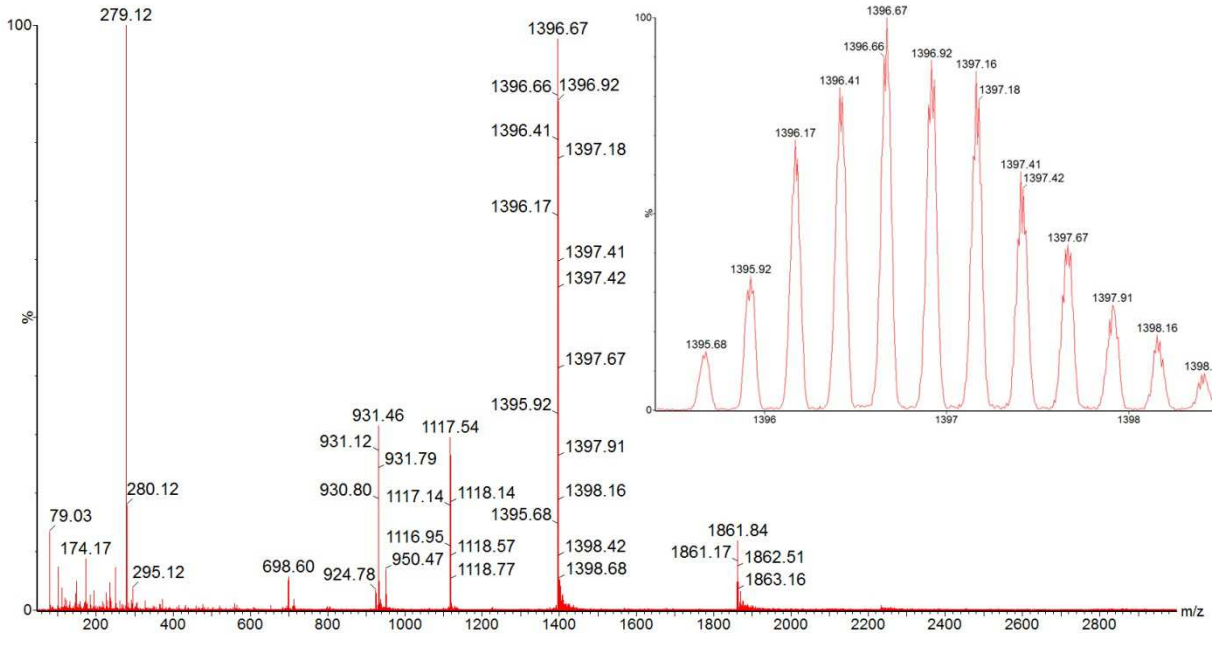


Figure 2.26. Mass spectrum of cyclic octamer **1c**$_8$ from the LC-MS analysis of a stirred library made from peptide **1c**. Calculated isotopic profile (species, abundance): 5580.35 (M+2, 100.0%), 5581.35 (M+3, 85.8%), 5579.35 (M+1, 77.4%); m/z calculated: 1861.12 $[(M+2)+3H]^{3+}$, 1396.08 $[(M+2)+4H]^{4+}$, 1117.07 $[(M+1)+5H]^{5+}$; m/z observed: 1861.84 $[(M+2)+3H]^{3+}$, 1396.67 $[(M+2)+4H]^{4+}$, 1117.54 $[(M+2)+5H]^{5+}$. Inset: isotopic profile at m/z 1396.67.

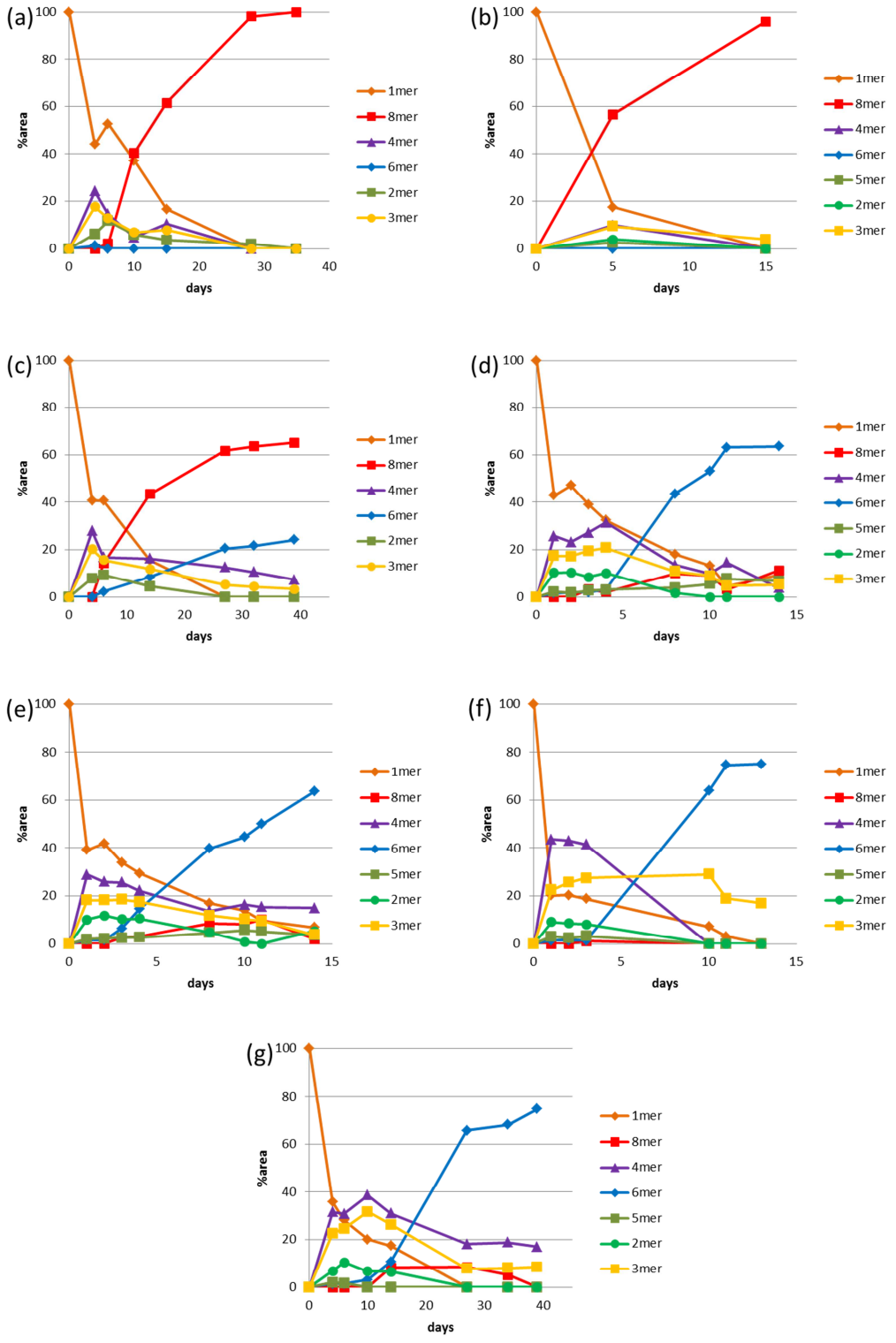


Figure 2.27. Evolution of the product distribution of DCLs (3.8 mM dithiol building block **1c** in 50 mM borate buffer pH 8.2) made from building block **1c** at different concentration of TFE. (a) 10%; (b) 13%; (c) 15%. In graphs (a-c) the predominant replicator is the octamer; (d) 17%; (e) 20%; (f) 25%; (g) 30%. In graphs (d-g) the predominant replicator is the hexamer.

Kinetic simulations of a system of two replicators competing for a common building block

Simulations were performed using the Berkeley Madonna software package, using the following model. Herein R1 and R2 are the two replicators, A is the common resource and k1 and k2 are the rate constants for the replication reactions of R1 and R2, respectively. The ratio of k1 to k2 was varied to produce the data shown in Figure 2.3c.

```
{Top model}


  {Reservoirs}
  d/dt (R1) = + dR1dt
    INIT R1 = 1e-9
  d/dt (R2) = + dR2dt
    INIT R2 = 1e-9
  d/dt (A) = + dAdt
    INIT A = 0.01

    {Flows}
    dR1dt = k1*A*R1
    dR2dt = k2*A*R2
    dAdt = -k1*A*R1-k2*A*R2

    {Functions}
    k1 = 1
    k2 = 2
{Globals}
{End Globals}
```

2.6. References

[1] Pross, A., (2012) *What is Life? How Chemistry becomes Biology*, Oxford University Press.

[2] Earl, D. J.; Deem, M. W. Evolvability is a selectable trait. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 11531-11536 (2004).

[3] Kashtan, N.; Noor, E.; Alon, U. Varying environments can speed up evolution. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 13711-13716 (2007).

[4] Bennett, A. F.; Dao, K. M.; Lenski R. E. Rapid evolution in response to high-temperature selection. *Nature* **346**, 79–81 (1990).

[5] Malakoutikhah, M.; Peyralans, J. J.-P.; Colomb-Delsuc, M.; Fanlo-Virgos, H.; Stuart, M. C. A.; Otto, S. Uncovering the selection criteria for the emergence of multi-building-block replicators from dynamic combinatorial libraries. *J. Am. Chem. Soc.* **135**, 18406–18417 (2013).

[6] (a) Buck, M. *Q.* Trifluoroethanol and colleagues: cosolvents come of age. Recent studies with peptides and proteins. *Rev. Biophys.* **31**:3, 297-355 (1998); (b) Diaz, M.D.; Fioroni, M.; Burger, K.; Berger, S. Evidence of complete hydrophobic coating of bombesin by trifluoroethanol in aqueous solution: An NMR spectroscopic and molecular dynamics study. *Chem. Eur.* **8**, 1663–1669 (2002); (c) Fioroni, M.; Diaz, M.D.; Burger, K.; Berger, S. Solvation phenomena of a tetrapeptide in water/trifluoroethanol and water/ethanol mixtures: A diffusion NMR, interolecular NOE and molecular dynamics study. *J. Am. Chem. Soc.* **124**, 7737–7744 (2002); (d) Roccatano, D.; Colombo, G.; Fioroni, M.; Mark, A.E. Mechanism by which 2,2,2-trifluoroethanol/water mixtures stabilize secondary-structure formation in peptides: A molecular dynamics study. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 12179–12184 (2002).

[7] Colomb-Delsuc, M.; Mattia, E.; Sadownik, J. W.; Otto, S. Exponential self-replication enabled through a fibre elongation/breakage mechanism. *Nat. Comm.* **6**, 7427-7434 (2015).

[8] Roccatano, D.; Colombo, G.; Fioroni, M.; Mark, A. Mechanism by which 2,2,2-trifluoroethanol/water mixtures stabilize secondary-structure formation in peptides: A molecular dynamics study. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 12179-12184 (2002).

[9] The total UPLC peak area of the libraries was also monitored over time. We observed that the resulting values are independent of the composition of the library, indicating that we detect all library material and that the molar absorbivity of the building blocks are essentially independent of the macrocycle in which the building block resides. See Figure 2.2 for details.

[10] (a) Nitschke, J. R. Molecular networks come of age. *Nature* **462**, 736-738 (2009); (b) Ludlow, R. F.; Otto, S. Systems Chemistry. *Chem. Soc. Rev.* **37**, 101-108 (2008); (c) Ruiz-Mirazo, K.; Briones, C.; de la Escosura, A. Prebiotic Systems Chemistry: New perspectives for the origins of life. *Chem. Rev.* **114**, 285–366 (2014); (d) De la Escosura, A.; Briones, C.; Ruiz-Mirazo, K. The Systems perspective at the crossroads between chemistry and biology. *J. Theor. Biol.* **381**, 11-22 (2015).

[11] Carnall, J. M. A.; Waudby, C. A.; Belenguer, A. M.; Stuart, M. C. A.; Peyralans, J. J.-P.; Otto, S. Mechanosensitive self-replication driven by self-organization. *Science* **327**, 1502-1506 (2010).